(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 333 379 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
05.04.2006 Bulletin 2006/14

(51) Int Cl.:
G06F 11/14 (2006.01) *G06F 3/06* (2006.01)

(21) Application number: **03250749.3**

(22) Date of filing: **05.02.2003**

(54) **Emulated backup tape drive using data compression**

Emuliertes Bandlaufwerk zur Datensicherung mit Datenkompression

Emulation de dérouleur de bande pour sauvegarde de données avec compression de données

(84) Designated Contracting States:
DE GB

(30) Priority: **05.02.2002 US 72437**

(43) Date of publication of application:
06.08.2003 Bulletin 2003/32

(73) Proprietor: **QUANTUM CORPORATION**
Milpitas, CA 95035 (US)

(72) Inventor: **Bolt, Thomas B.**
Encinitas,
California 92024 (US)

(74) Representative: **Charig, Raymond Julian**
Eric Potter Clarkson LLP
Park View House
58 The Ropewalk
Nottingham NG1 5DD (GB)

(56) References cited:
EP-A- 0 869 439
EP-A- 1 164 491
WO-A-99/12098
US-B1- 6 341 329

EP 1 333 379 B1

Description

FIELD OF THE INVENTION

[0001] The present invention relates generally to back up data storage, and more specifically, to an emulated backup tape drive that stores non-compressed data during backup operations and then afterwards when the drive is idle, retrieves, compresses and then re-stores the data to reclaim space on the storage medium of the drive.

BACKGROUND

[0002] With the increasing popularity of Internet commerce and network centric computing, businesses and other entities are becoming more and more reliant on information. Protecting critical data from loss due to system crashes, virus attacks and the like is therefore of primary importance. A well designed data protection program will generally have the ability to (i) instantly re-store data in the event of a disaster to enabled continued computing operations; (ii) re-store data over an extended period of time (hours or days) without disrupting normal computing operations; and (iii) archive copies of data that are retrieved infrequently and with little urgency. Tape drives have long been a choice for storing archival back up data in information systems.

[0003] Historically many such tape drives have used data compression to maximize the amount of data that can be stored on the tape. Tape, however, is a relatively slow and inefficient storage medium. Consequently emulated "tape" drives that use arrays of hard drives have become more popular recently. These emulated tape drives often rely on data compression to enable the storage of more data. The problem with current emulated tape drives is that the data compression is performed "on the fly" during the backup. In other words, compression occurs in the critical path of the down loading of data, thereby impeding performance. The designers of emulated tape drive systems have therefore relied on expensive, high speed, hardware data compression solutions to achieve an acceptable level of performance. The use of slower, less expensive software compression algorithms have not been a viable option in the past because of a lack of acceptable performance.

[0004] US 6 341 329 discloses a backup system emulating a tape drive or a hard disk. EP 1 164 491 discloses compression of backup data scheduled as a function of utilised disk space.

[0005] An emulated backup tape drive that stores non-compressed data during backup operations and then afterwards when the drive is idle, retrieves, compresses and then re-stores the data to reclaim space on the storage medium of the drive is therefore needed.

SUMMARY

[0006] To achieve the foregoing, and in accordance with the purpose of the present invention as defined by the claims, a back up storage device is disclosed that stores non-compressed data during backup operations and then afterwards when the device is idle, retrieves, compresses and re-stores the data to reclaim space on the storage medium of the device. During operation, a duty cycle having a backup window period and an idle period is defined. When back ups occur during the window, data is down-loaded and stored on the device in non-compressed form. Later during the idle period of the duty cycle, the non-compressed data is retrieved, compressed and re-stored to reclaim space on the storage medium of the device. Since the compression occurs when the back up device is idle, the rate at which data is backed up is not adversely effected in any way. Thus a low cost software data compression algorithm may be used. In one embodiment of the invention, the back up storage device is an emulated tape drive that uses an array of hard drives for the storage medium. In other embodiments, any type of storage medium can be used such as tape or semiconductor memory chips for example

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] The invention, together with further advantages thereof, may best be understood by reference to the following description taken in conjunction with the accompanying drawings in which:

Figure 1 is block diagram of an exemplary information infrastructure in which the emulated backup tape drive of the present invention may be used.

Figures 2A, 2B and 2C are diagrams of the emulated tape drive of the present invention.

Figure 3 is block diagram of a controller of the emulated tape drive of the present invention.

Figure 4 is flow diagram illustrating a backup and compression duty cycle of the emulated tape drive of the present invention.

DESCRIPTION

[0008] Referring to Figure 1, a block diagram of an exemplary information infrastructure in which the emulated backup tape drive of the present invention may be used is shown. The information infrastructure 10 includes a plurality of clients 12 and a plurality of servers 14 coupled together by a client network 16, a primary storage location 18, and one or more emulated tape drives 20 coupled together by a network connection 22. The clients 12 can be any type of client such as but not limited to a personal computer, a "thin" client, a personal digital as-

sistant, a web enabled appliance or a cell phone. The servers 14 may also be of any variety such as those based on the Unix, Linux, or the Microsoft Windows operating systems or a combination thereof. Likewise, the client network 16 can be any type of network including but not limited to the Internet, a corporate intranet, a wide area network, a local area network, a wireless network, or any combination thereof. The primary storage location 18 can be arranged in a number of different types of configurations, such as a storage array network (SAN), or network attached storage (NAS), or direct attached storage . In other embodiments, the primary storage location 18 can reside in the chassis/cabinet of the servers 14, stand alone storage devices, or a combination thereof. The connection 22 can be either direct (such as parallel SCSI or IDE) or a network topology such as fibre channel, Ethernet (fast, gigabit, or 10 gigabit) for example. Also when multiple emulated tape devices 20 are used, they may be daisy-chained together to provide more backup capacity.

[0009] Referring to Figure 2A, a perspective view of an emulated tape drive 20 is shown. The emulated tape drive 20 includes a chassis 30, a power supply 32, a pair of fans 34, and an array of hard drives 36, all housed in the chassis 30. A user interface panel 38 is located at the front of the chassis 30. A back plate 40 is provided at the rear portion of the chassis 30.

[0010] Referring to Figure 2B, an exploded perspective view of the hard drives 36 according to one embodiment of the present invention is shown. In this embodiment, the hard drives 36 are configured in a pair of left and right rails (42L and 42R) respectively. Within each rail 42, five (5) columns of three (3) hard drives 36 are arranged in disk packs 44 respectively. It should be noted that this embodiment is only exemplary. According to various other embodiments of the present invention, the number of hard drives 36 may be arranged in any number of rails 42 (rows) and the number of hard drives 36 (columns) per disk pack 44 may vary.

[0011] Referring to Figure 2C, a view of the back plate 40 of the chassis 30 is shown. The back plate 40 includes vents for the fans 34 and a number of input/output ports 46. The input/output ports 46 are provided to connect a controller 48 (not visible because it is internal to the chassis 30) to the primary storage location 18 through the network connection 22. For more details on the features and operation of the emulated tape drive 20, see co-pending EP application entitled "Storage System Utilizing An Active Subset of Drives During Data Storage And Retrieval Operations" by Thomas B. Bolt and Kevin C. Daly, attorney docket no. Q02-1037.US1, and published as EP 1 336 921.

[0012] Referring to Figure 3, a block diagram of one embodiment of the controller 48 configured for the emulated tape drive 20 illustrated in Figures 2A-2C is shown. With this embodiment, the controller 48 includes a micro-controller 50, such as a microprocessor, configured to communicate with the hard drives 36 of the disk packs 44 through a USB controller 52, a USB hub 54 and bridge circuit 56. For the sake of simplicity, these components are shown for only one disk pack 44. The remaining four disk packs 44 of the right rail 42R and all of the disk packs 44 of the left rail 42L communicate with the micro-controller 50 in a similar arrangement. In situations where the network connection 22 is fiber channel, the micro-controller 50 is connected to the primary storage location 18 through an optical transceiver 58 and a fiber channel controller 60. Alternatively, when the network connection 22 is a Giga-bit Fast ethernet connection, the micro-controller 50 is connected to the primary storage location 18 through an ethernet transceiver 62 and an ethernet controller 64. It should be pointed out that these two connections are merely illustrative. In various embodiments of the present invention, multiple fiber channel ports and/or multiple ethernet channel ports, either alone or in any combination, can be provided. Alternate interfaces such as parallel SCSI (Small Computer System Interface) may also be substituted or used in conjunction with fiber channel or Ethernet. Additionally, alternate internal interconnect technologies such as fibre channel or parallel SCSI could be used instead of USB.

[0013] A system memory 66 and a non-volatile memory 68 are also coupled to the micro-controller 50. In one embodiment of the invention, the system memory 66 is RAM and the non-volatile memory is Flash. The non-volatile memory is used for storing the micro-code used to program the micro-controller 50 as well as the compression/decompression software algorithms which are used by the emulated tape drive 20. It again should be noted that the circuit components of this diagram are merely illustrative of one embodiment of the present invention. Other embodiments would be readily apparent to those skilled in the art. For example, in embodiments with either more or fewer rails 42 and disk packs 44, additional or fewer USB controllers 52 and USB hubs 54 would be required. Also the interface hardware between the input/output ports 46 would be different if other type of networking protocols besides fiber channel or ethernet were used.

[0014] Referring to Figure 4, a flow diagram 80 illustrating the operation of the emulated tape drive 20 is shown. Initially a system administrator or user of the information infrastructure 10 defines a duty cycle (step 82) for the emulated tape drive 20. The duty cycle includes a backup window period and an idle period. Usually the backup window period is scheduled at a set time each day or at some other fixed time interval when the emulated tape drive 20 is at its lowest utilization. When backups are not occurring, the emulated tape drive 20 is idle.

[0015] According to one embodiment, when the duty cycle starts (step 84), the backup of data begins (step 86). The data is downloaded from the primary storage location 18 through the input/output ports 46, the micro-controller 50, the appropriate USB controller 52, USB hub 54, and bridge circuit 56 and stored (step 88) in non-compressed form on one of the hard drives 36. When

the backup window period expires, the idle period begins (step 92). During the idle period, the non-compressed data stored on the hard drives 36 is retrieved (step 94) and provided to the micro-controller 50 through the bridge circuit 56, USB hub 54, and USB controller 52. The micro-controller 50 compresses the data (step 96) and then re-stores it on the hard drives 36 (step 98) to reclaim space on the hard drives 36. When the idle period is over and next backup window begins, a new duty cycle begins (step 84) and the aforementioned steps are repeated. In various embodiments, any one of a variety of software compression algorithms may be used, such as a zip; a gnuzip; a bzip; a b2zip; a Lempil Ziv; and a LZS (Lempil Ziv Stac). Alternately, other compression algorithms can be used.

[0016] When compressed data on the emulated tape drive 20 is needed, it is retrieved and provided to the micro-controller 50. The data is decompressed using the software algorithms stored in the Flash memory 68 and provided to the primary storage location 18 through appropriate input/output port 46. Since compression algorithms are typically asymmetric, data decompression is not nearly as computationally intensive as compression, the performance of the emulated tape drive 20 during data retrieval is not significantly degraded using a software solution.

[0017] The present invention thus provides an emulated tape device used for the backup of archival data that uses a software based data compression algorithm. Since the compression occurs when the emulated tape drive 20 is idle, the rate at which data is stored during a backup operation is not adversely effected in anyway.

[0018] It should be noted that the duty cycle could occur any time there is a detection of inactivity on the primary data interface 46. After some time period of inactivity (for example, 20 minutes), the system could begin retrieving and compressing data. This process can be interrupted at any point if activity is detected on the primary data interface 46. It is acceptable for data to be partially compressed, and it is possible to restart the compression from the point at which it was previously suspended.

[0019] Although the foregoing invention has been described in some detail for purposes of clarity of understanding, it will be apparent that certain changes and modifications may be practiced within the scope of the appended claims. For instance, the present invention can be ready practice with any type of read-write storage medium such as magnetic tape, silicon memory chips, devices such as SRAM, DRAM, Flash, EPROM, EUVPROM. EEPROM, etc. The described embodiments should therefore be taken as illustrative and not restrictive, and the invention should not be limited to the details given herein but should be defined by the following claims and their full scope.

Claims

1. A method, comprising:

   defining a duty cycle for the downloading of data to a backup storage device, the duty cycle having a backup window period and an idle period during which no backup occurs;
   receiving data during the backup window period;
   storing the data on the backup storage device during the backup window period;
   retrieving the data stored on the backup storage device during the idle period after the backup window period;
   compressing the data retrieved from the backup storage device during the idle period; and
   re-storing the data compressed during the idle period in compressed form on the backup storage device during the idle period to reclaim space on the storage device.

2. The method of claim 1, wherein the compression of data is performed using a software data compression algorithm.

3. The method of claim 2, wherein the software data compression algorithm includes one of the following types of algorithms: a zip; a gnuzip; a bzip; a b2zip; a Lempil Ziv; and a LZS (Lempil Ziv Stac).

4. The method of claim 1, further comprising successively repeating the receiving and storing of data during the backup window periods and retrieving, compressing and storing compressed data on the backup storage device during successive duty cycles respectively.

5. The method of claim 1, wherein the backup storage device is an emulated tape drive containing an array of hard drives.

6. The method of claim 1, wherein the data is downloaded over a network from a primary storage location.

7. The method of claim 6, wherein the data is downloaded over a fiber-channel connection between the primary storage location and the backup storage device.

8. The method of claim 6, wherein the data is downloaded over an ethemet connection between the primary storage location and the backup storage device.

9. The method of claim 6, wherein the primary storage location and the backup storage device are part of a storage array network.

10. The method of claim 6, wherein the primary storage location and the backup storage device are part of a network attached storage configuration.

11. The method of claim 1, wherein the backup storage device is directly electrically connected to a server.

12. An apparatus comprising:

a backup storage device comprising:

an input/output port;
an array of hard drives configured as backup storage; and
a controller configured to download data received from the input/output port to the array of hard drives during a backup period and then reclaim storage space on the array of hard drives during an idle period following the backup period and during which no backup occurs, by retrieving the data stored on the array of hard drives, compressing the retrieved data, and then re-storing the compressed data on the array of hard drives.

13. The apparatus of claim 12, wherein the controller is further configured to execute a software algorithm to compress the retrieved data.

14. The apparatus of claim 13, wherein the software algorithm includes one of the following types of algorithms a zip; a gnuzip; a bzip; a b2zip; a Lempil Ziv; and a LZS (Lempil Ziv Stac).

15. The apparatus of claim 13, wherein the software algorithm is stored in a memory associated with the controller.

16. The apparatus of claim 12, further comprising a fiber channel controller coupled between the controller and the input/output port which comprises an optical transceiver.

17. The apparatus of claim 12, further comprising an ethernet controller coupled between the controller and the input/output port which comprises an ethernet transceiver.

18. The apparatus of claim 12, wherein the array of hard drives configured as backup storage further comprises a network hub and bridge circuit coupled between the array of hard drives and the controller.

19. The apparatus of claim 12, further comprising:

a primary storage location coupled to the backup storage device through a network connection.

20. The apparatus of claim 19, wherein the network connection is one of the following types of network connections: fiber channel or ethernet.

21. The apparatus of claim 19, wherein the primary storage location and the backup storage device are arranged in one of the following: a storage attached network or network attached storage configuration.

22. The apparatus of claim 19, further comprising a plurality of clients and servers coupled to the primary storage location through a client network.

Patentansprüche

1. Verfahren, mit:

Festlegen eines Arbeitszyklus für das Herunterladen von Daten in eine Sicherungsspeichervorrichtung, wobei der Arbeitszyklus ein Sicherheitsfensterzeitintervall und ein Ruhezeitintervall, währenddessen keine Sicherung stattfindet, aufweist;
Empfangen von Daten während des Sicherungsfensterzeitintervalls;
Speichern der Daten auf der Sicherungsspeichervorrichtung während des Sicherungsfensterzeitintarvalls;
Wiedergewinnen der auf Sicherungsspeichervorrichtung gespeicherten Daten während der Ruhezeitspanne nach der sicherungsfensterzeitspanne;
Komprimieren der von der sicherungsspeichervorrichtung wiedergewonnenen Daten während des Ruhezeitintervalls; und
Wieder-Speichern der während des Ruhezeitintervalls komprimierten Daten in komprimierter Form auf der Sicherungsspeichervorrichtung während der Ruhezeitspanne, um erneut Platz auf der Speichervorrichtung zu beanspruchen.

2. Verfahren gemäß Anspruch 1, bei dem die Komprimierung von Daten mit einem Softwaredatenkomprizierungsalgorithmus durchgeführt Wird.

3. Verfahren gemäß Anspruch 2, bei dem der Softwaredatenkomprimierungsalgorithmus eine der folgenden Arten von Algorithmen aufweist: ein zip; ein gnuzip; ein bzip; ein b2zip: ein Lempil Ziv; und LZS (Lempil Ziv Stac).

4. Verfahren gemäß Anspruch 1, ferner mit sukzessivem Wiederholen des Empfangens und Speicherns von Daten während der Sicherungsfensterzeitintervalle und des Wiedergewinnens, Komprimierens bzw. Speichern von komprimierten Daten auf der Sicherungsspeichervorrichtung während aufeinan-

derfolgender Arbeitszyklen.

5. Verfahren gemäß Anspruch 1, bei dem die Sicherungsspeichervorrichtung ein emuliertes Bandlaufwerk ist, das ein Array von Festplatten enthält.

6. Verfahren gemäß Anspruch 1, bei dem die Daten über ein Netzwerk von einer primären Speicherstelle heruntergeladen werden.

7. Verfahren gemäß Anspruch 8, bei dem die Daten über eine Faserkartalverbindung zwischen der primären Speicherstelle und der sicherungsspeichervorrichtung heruntergeladen werden.

8. Verfahren gemäß Anspruch 6, bei dem die Daten über eine Ethernet-Verbindung zwischen der primären Speicherstelle und der Sicherungsspeichervorrichtung heruntergeladen werden.

9. Verfahren gemäß Anspruch 6, bei dem die primäre. Speicherstelle und die Sicherungsspeichervorrichtung Teil eines Speicherarraynetzwerks sind.

10. Verfahren gemäß Anspruch 6, bei dem die primäre speicherstelle und die Sicherungsspeiahervorriahtung Teil einer mit einem Netzwerk verbundenen Speicherkonfiguration sind.

11. Verfahren gemäß Anspruch 1, bei dem die Sicherungsspeichervorrichtung mit einem Server direkt elektrisch verbunden ist.

12. Vorrichtung, mit:

einer Sicherungsspeichervorrichtung mit:

einem Eingabe/Ausgabe-Port;
einem Array von Festplatten, die als Sicherungsspeicher konfiguriert sind; und
einem controller, der konfiguriert ist, um von dem Eingabe/Ausgabe-Port empfangene Daten in das Array von Festplatten während eines Sicherungszeitintervalls herunterzuladen und um dann, während eine Ruhezeitintervalls nach dem sicherungszeitintervall, währenddessen keine Sicherung stattfindet, Speicherplatz auf dem Array von Festplatten zu beanspruchen, wobei die auf dem Array von Festplatten gespeicherten Daten wiedergewonnen, die wiedergewonnenen Daten komprimiert und die komprimierten Daten auf dem Array von Festplatten erneut gespeichert werden.

13. Vorrichtung gemäß Anspruch 12, bei der der Controller ferner konfiguriert ist, um einen Softwarealgorithmus auszuführen, um die wiedergewonnenen Daten zu komprimieren.

14. Vorrichtung gemäß Anspruch 13, bei der der Softwarealgorithmus eine der folgenden Arten von Algorithmus aufweist: ein zip; ein gnuzip; ein bzip; ein b2zip; ein Lempil Ziv; und ein LZS (Lsrngil Ziv Stac) .

15. Vorrichtung gemäß Anspruch 13, bei der der Softwarealgorithmus in einem dem Controller zugeordneten Speicher gespeichert ist.

16. Vorrichtung gemäß Anspruch 12, ferner mit einem Faserkanalcontroller, der zwischen dem Controller und dem Eingabe/Ausgabe-Port gekoppelt ist, der einen optischen Transceiver umfasst.

17. Vorrichtung gemäß Anspruch 12, ferner mit einem Ethernet-controller, der zwischen dem Controller und dem Eingabe/Ausgabe-Port gekoppelt ist, der einen Ethernet-Transceiver umfasst.

18. Vorrichtung gemäß Anspruch 12, bei der das Array von Festplatten, das als Sicherungsspeicher konfiguriert ist, ferner einen Netzwerkhub und eine Brükkenschaltung, die zwischen dem Array von Festplatten und dem Controller gekoppelt ist, umfasst.

19. Vorrichtung gemäß Anspruch 12, ferner mit:

einer primären Speicherstelle, die mit der Sicherungsspeichervorrichtung durch eine Netzwerkverbindung gekoppelt ist.

20. Vorrichtung gemäß Anspruch 19, bei der die Netzwerkverbindung eine der folgenden Arten von Netzwerkverbindungen ist: Faserkanal oder Ethernet.

21. Vorrichtung gemäß Anspruch 19, bei der die primäre Speicherstelle und die sicherungsspeichervorrichtung in einem der folgenden angeordnet sind: einem Speicher-Netzwerk oder einer mit einem Netzwerk verbundenen Speicherkonfiguration.

22. Vorrichtung gemäß Anspruch 19, ferner mit einer Mehrzahl von Clienten und Servern, die mit der primären speicherstelle durch ein Client-Netzwerk gekoppelt sind.

Revendications

1. Procédé comprenant les étapes consistant à :

définir un cycle actif pour le téléchargement de données sur un dispositif de stockage pour sauvegarde, le cycle actif comprenant une période de fenêtre de sauvegarde et une période de repos durant laquelle aucune sauvegarde ne se

produit ;
recevoir des données au cours de la période de fenêtre de sauvegarde;
enregistrer les données sur le dispositif de stockage pour sauvegarde au cours de la période de fenêtre de sauvegarde ;
retrouver les données enregistrées sur le dispositif de stockage pour sauvegarde au cours de la période de repos après la période de fenêtre de sauvegarde ;
compresser les données retrouvées à partir du dispositif de stockage pour sauvegarde au cours de la période de repos ; et
réenregistrer les données compressées au cours de la période de repos sous une forme compressée sur le dispositif de stockage pour sauvegarde au cours de la période de repos de manière à gagner de la place sur le dispositif de stockage.

2. Procédé selon la revendication 1, dans lequel la compression de données est exécutée en utilisant un algorithme de compression de données informatique.

3. Procédé selon la revendication 2, dans lequel l'algorithme de compression de données informatique comprend l'un des types d'algorithmes suivants : un algorithme zip ; un algorithme gnuzip ; un algorithme bzip ; un algorithme b2zip ; un algorithme Lempil Ziv ; et un algorithme LZS (Lempil Ziv Stac).

4. Procédé selon la revendication 1, comprenant en outre les étapes consistant à répéter successivement la réception et le stockage de données au cours des périodes de fenêtre de sauvegarde et consistant à retrouver, compresser et enregistrer les données compressées sur le dispositif de stockage pour sauvegarde au cours de cycles actifs successifs, respectivement.

5. Procédé selon la revendication 1, dans lequel le dispositif de stockage pour sauvegarde est un dérouleur de bande émulé pour sauvegarde de données contenant une matrice de disques durs.

6. Procédé selon la revendication 1, dans lequel les données sont téléchargées sur un réseau à partir d'un emplacement de stockage primaire.

7. Procédé selon la revendication 6, dans lequel les données sont téléchargées sur une connexion « Fiber Channel » entre l'emplacement de stockage primaire et le dispositif de stockage pour sauvegarde.

8. Procédé selon la revendication 6, dans lequel les données sont téléchargées sur une connexion Ethernet entre l'emplacement de stockage primaire et le dispositif de stockage pour sauvegarde.

9. Procédé selon la revendication 6, dans lequel l'emplacement de stockage primaire et le dispositif de stockage pour sauvegarde font partie d'un réseau à matrice de stockage.

10. Procédé selon la revendication 6, dans lequel l'emplacement de stockage primaire et le dispositif de stockage pour sauvegarde font partie d'une configuration de stockage attachée à un réseau.

11. Procédé selon la revendication 1, dans lequel le dispositif de stockage pour sauvegarde est directement connecté électriquement à un serveur.

12. Appareil comprenant :

un dispositif de stockage pour sauvegarde comprenant :

un port d'entrée/sortie :
une matrice de disques durs configurée en tant qu'une matrice de stockage de données de sauvegarde ; et
un contrôleur configuré pour télécharger des données reçues depuis le port d'entrée/sortie dans la matrice de disques durs au cours d'une période de sauvegarde et pour gagner de la place pour le stockage sur la matrice de disques durs au cours d'une période de repos qui suit la période de sauvegarde et pendant laquelle aucune sauvegarde ne se produit, en retrouvant les données enregistrées sur la matrice de disques durs, en compressant les données retrouvées, et en réenregistrant ensuite les données compressées sur la matrice de disques durs.

13. Appareil selon la revendication 12, dans lequel le contrôleur est configuré par ailleurs pour exécuter un algorithme informatique dans le but de compresser les données retrouvées.

14. Appareil selon la revendication 13, dans lequel l'algorithme de compression de données informatique comprend l'un des types d'algorithmes suivants : un algorithme zip ; un algorithme gnuzip ; un algorithme bzip ; un algorithme b2zip ; un algorithme Lempil Ziv ; et un algorithme LZS (Lempil Ziv Stac).

15. Appareil selon la revendication 13, dans lequel l'algorithme informatique est enregistré dans une mémoire associée au contrôleur.

16. Appareil selon la revendication 12, comprenant en

outre un contrôleur « Fiber Channel » couplé entre le contrôleur et le port d'entrée/sortie qui comprend un appareil émetteur-récepteur optique.

17. Appareil selon la revendication 12, comprenant en outre un contrôleur Ethernet couplé entre le contrôleur et le port d'entrée/sortie qui comprend un appareil émetteur-récepteur Ethernet.

18. Appareil selon la revendication 12, dans lequel la matrice de disques durs configurée en tant qu'une matrice de stockage de données de sauvegarde comprend en outre un concentrateur de réseau et un circuit en pont qui sont couplés entre la matrice de disques durs et le contrôleur.

19. Appareil selon la revendication 12, comprenant en outre :

un emplacement de stockage primaire couplé au dispositif de stockage pour sauvegarde par le biais d'une connexion réseau.

20. Appareil selon la revendication 19, dans lequel la connexion réseau est une connexion réseau de l'un des types suivants : « fiber channel » ou ethernet.

21. Appareil selon la revendication 19, dans lequel l'emplacement de stockage primaire et le dispositif de stockage pour sauvegarde sont agencés selon l'un de ce qui suit : un réseau attaché au dispositif de stockage ou une configuration de stockage attachée au réseau.

22. Appareil selon la revendication 19, comprenant en outre une pluralité de clients et de serveurs couplés à l'emplacement de stockage primaire par le biais d'un réseau client.

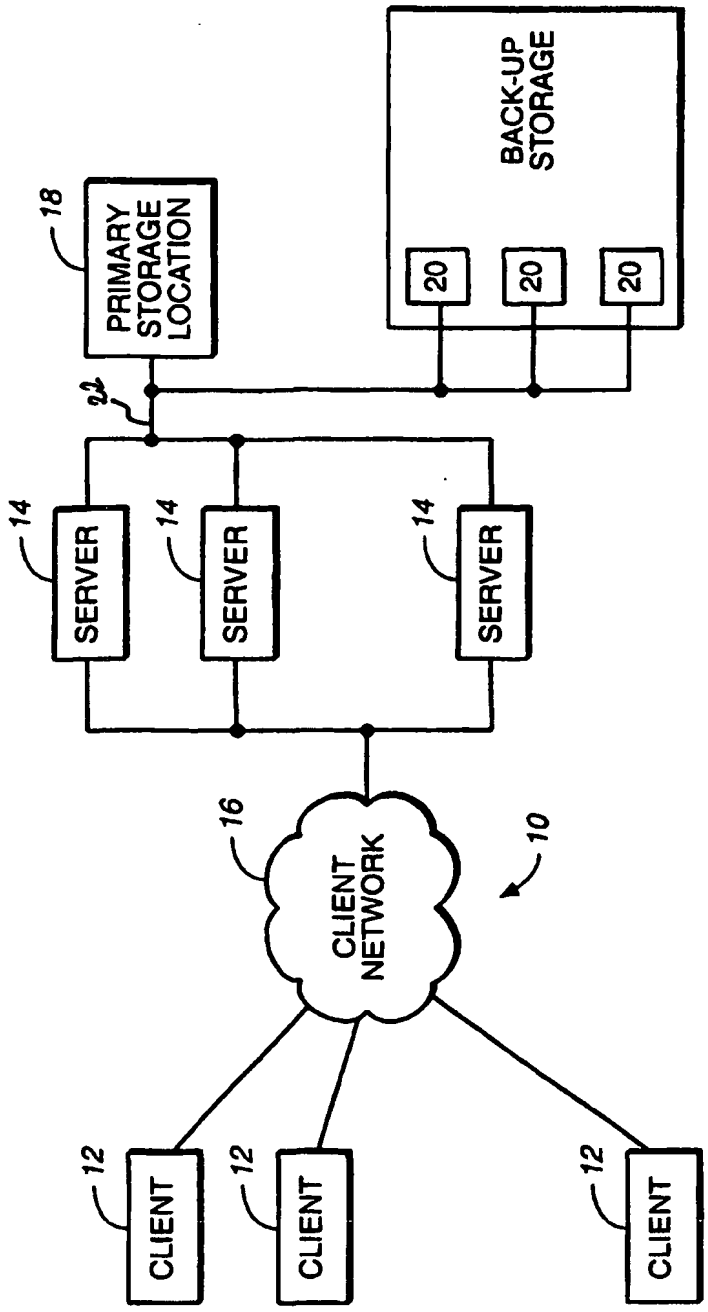
BACK-UP STORAGE
20
20
20
18
PRIMARY STORAGE LOCATION
22
14
SERVER
14
SERVER
14
SERVER
16
CLIENT NETWORK
10
12
CLIENT
12
CLIENT
12
CLIENT

FIG._1

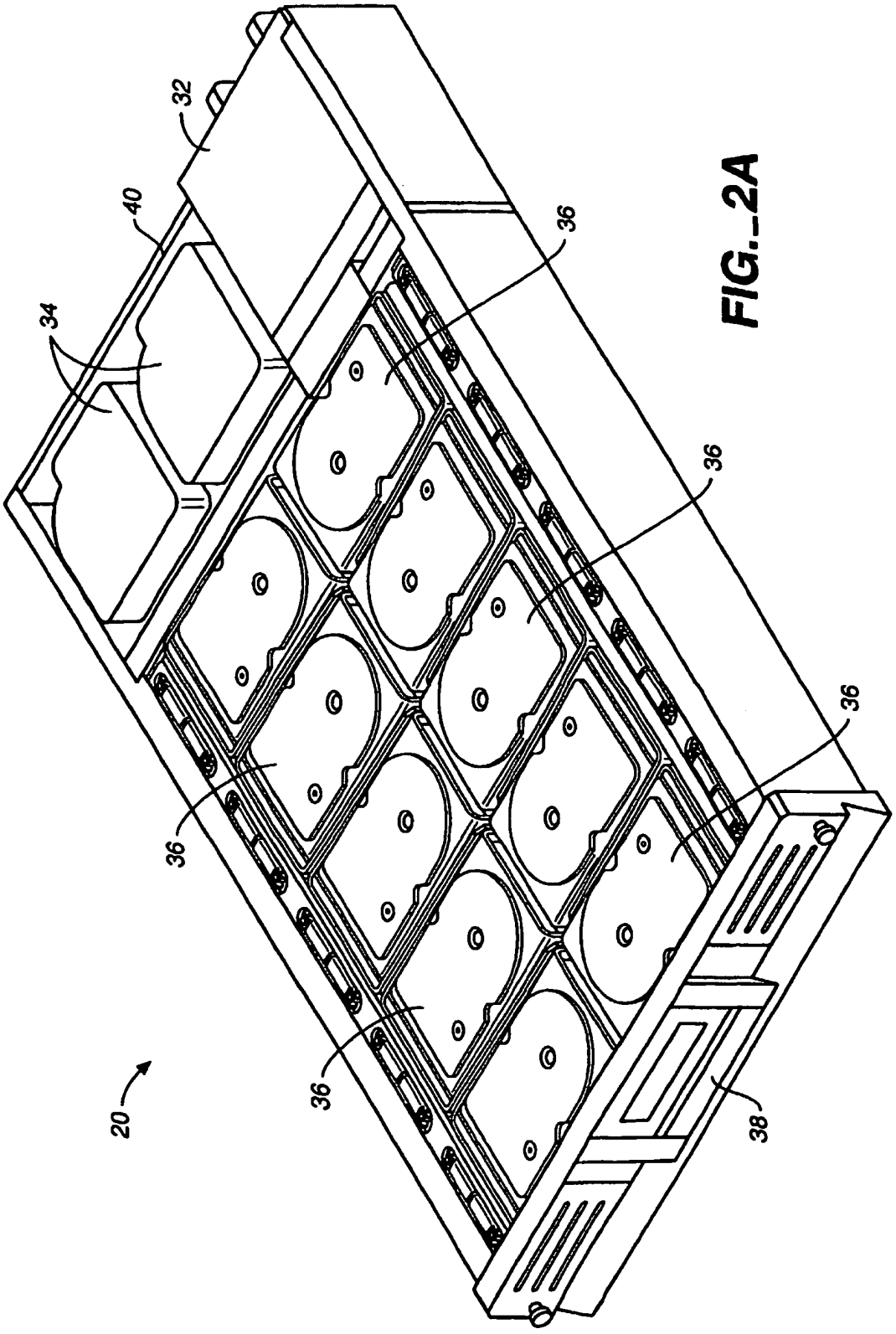

FIG._2A

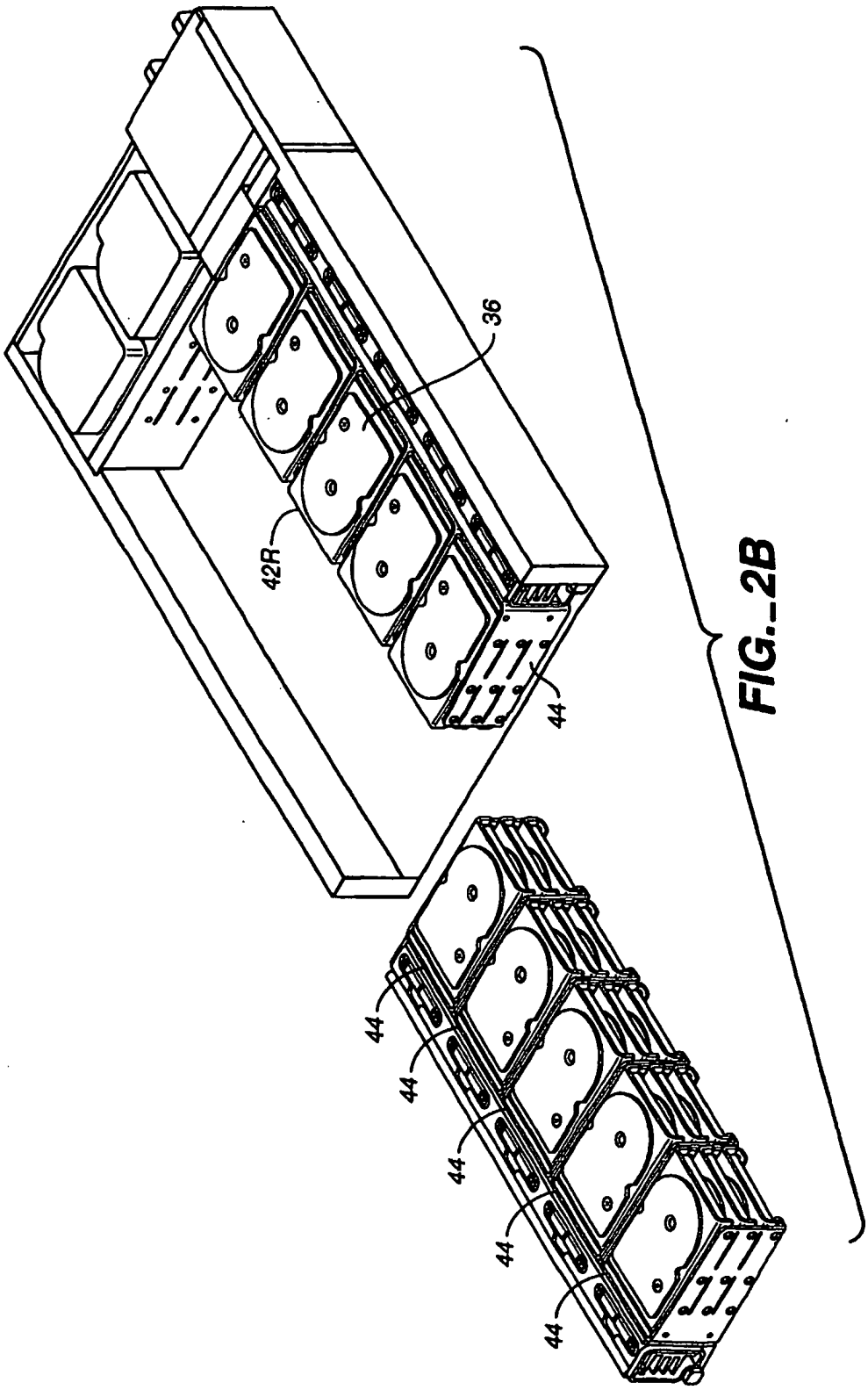

FIG._2B

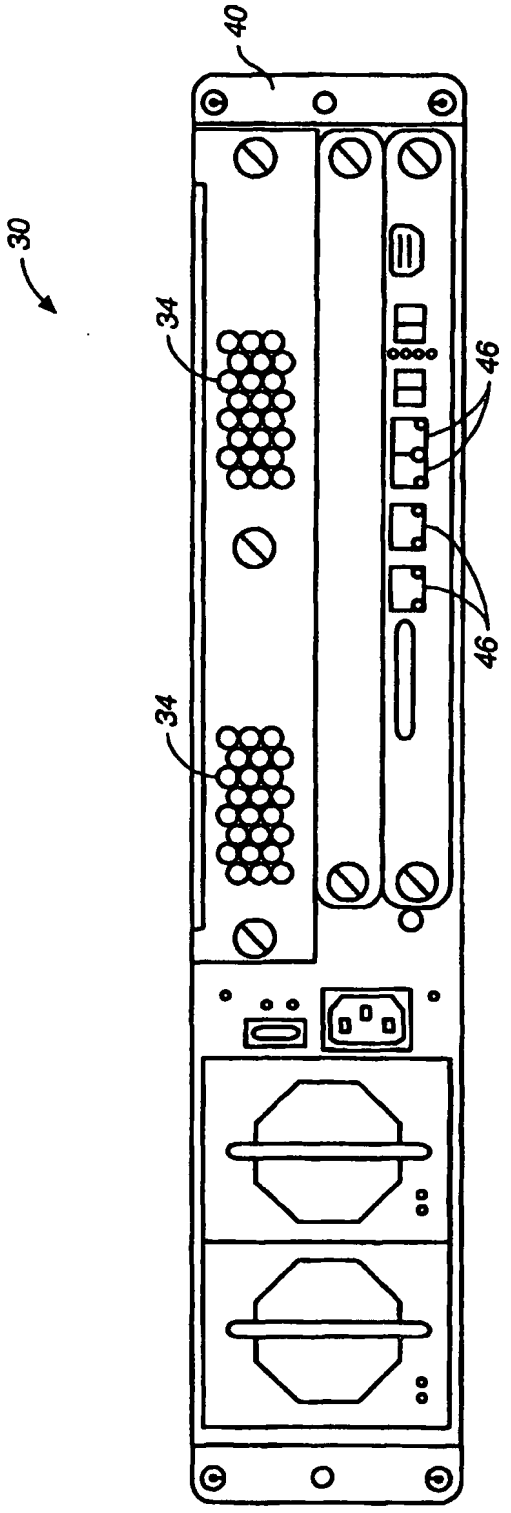

FIG._2C

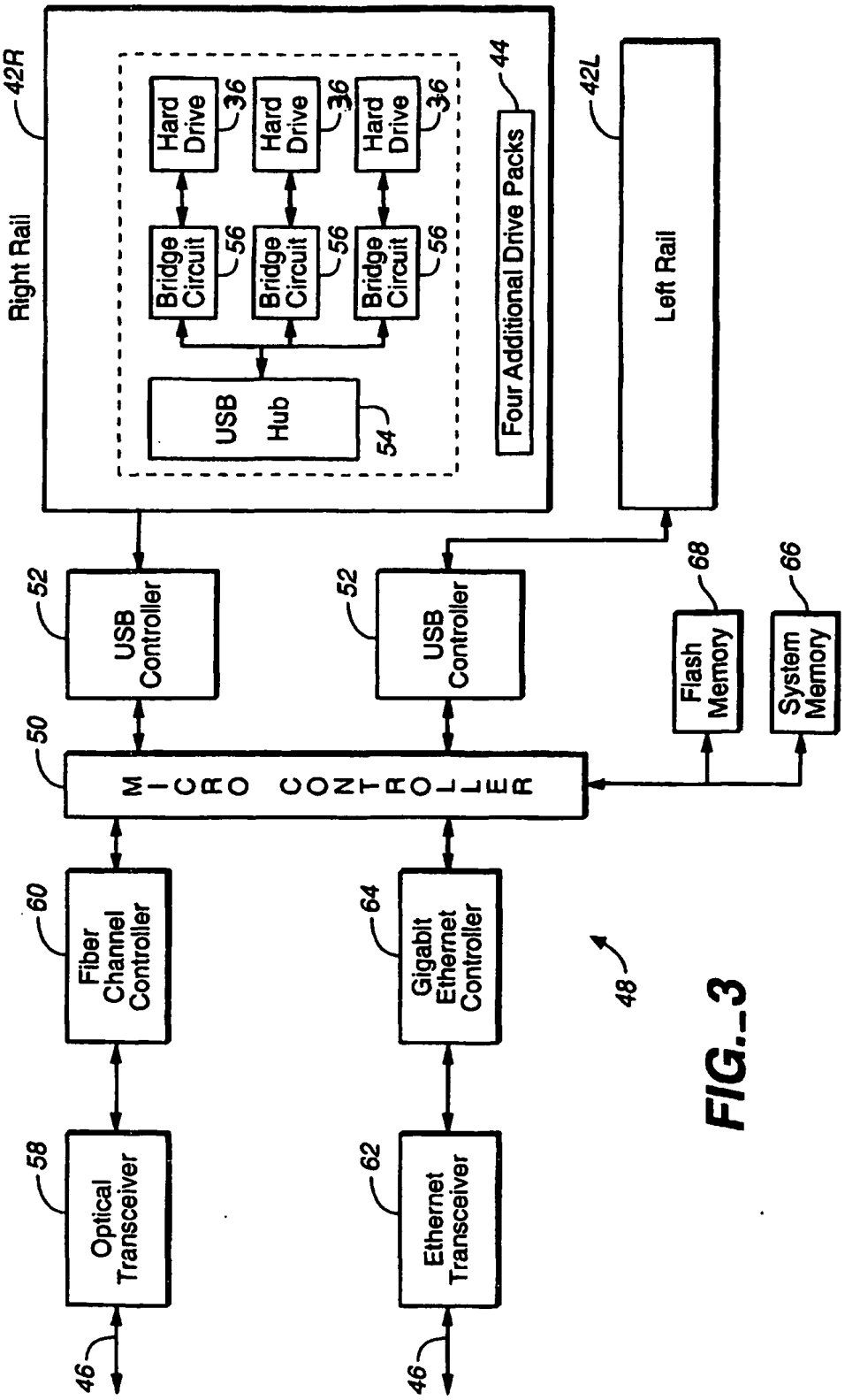
42R
Right Rail
Hard Drive
36
Bridge Circuit
56
USB Hub
54
44
Four Additional Drive Packs
42L
Left Rail
52
USB Controller
50
M-CRO CONTROLLER
68
Flash Memory
66
System Memory
60
Fiber Channel Controller
64
Gigabit Ethernet Controller
58
Optical Transceiver
62
Ethernet Transceiver
46
48

FIG._3

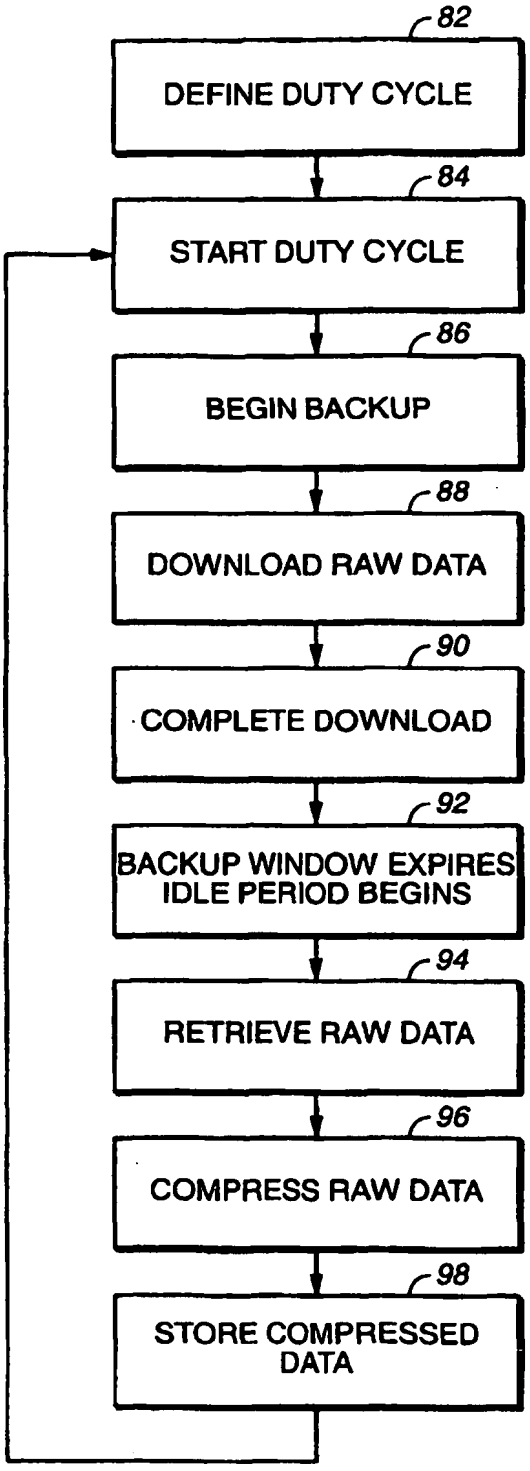
82
DEFINE DUTY CYCLE
84
START DUTY CYCLE
86
BEGIN BACKUP
88
DOWNLOAD RAW DATA
90
COMPLETE DOWNLOAD
92
BACKUP WINDOW EXPIRES
IDLE PERIOD BEGINS
94
RETRIEVE RAW DATA
96
COMPRESS RAW DATA
98
STORE COMPRESSED
DATA

FIG._4